IMAGE RECOGNIZING DEVICE AND METHOD THEREFOR

Patent Number:
Publication date: 1999-08-10
Inventor(s): MIHARA NORIO; DOI MIWAKO
Applicant(s): TOKYO SHIBAURA ELECTRIC CO
Requested Patent: JP11219421
Application Number: JP19980019871 19980130
Priority Number(s): JP19980019871 19980130
IPC Classification: G06T1/00; G10L3/00
EC Classification: G06K9/00D, G06K9/00F2, G06K9/00G
Equivalents: JP3688879B2, US2002126879, US6504944

Abstract

PROBLEM TO BE SOLVED: To provide an image recognizing device capable of quickly and highly accurately recognizing the shape or movement of the mouth and lip of a human. SOLUTION: This device is provided with an image acquiring part 1 for acquiring a depth map stream of an object, an oral cavity part extraction part 2 for extracting an oral cavity part from the stream acquired by the acquiring part 1 and an image recognition part for recognizing at least one of the shape and movement of mouth and lip based on the depth map stream of the oral cavity part extracted by the extraction part 2. Since a necessary part is extracted from the depth map of the object and recognition processing is executed based on the distance image of the extracted part, the shape or movement of a human face or mouth and lip can be quickly and highly accurately recognized.

(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号

特開平11－219421

(43)公開日　平成11年(1999)８月10日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｇ０６Ｔ　1/00		Ｃ０６Ｆ　15/62	３８０
Ｇ１０Ｌ　3/00	５１３	Ｃ１０Ｌ　3/00	５１３Ｚ
	５７１		５７１Ｃ
			５７１Ｋ

審査請求　未請求　請求項の数15　ＯＬ　（全 20 頁）

(21)出願番号　特願平10－19871

(22)出願日　平成10年(1998)１月30日

(71)出願人　000003078
株式会社東芝
神奈川県川崎市幸区堀川町72番地

(72)発明者　三原　功雄
神奈川県川崎市幸区小向東芝町１番地　株式会社東芝研究開発センター内

(72)発明者　土井　美和子
神奈川県川崎市幸区小向東芝町１番地　株式会社東芝研究開発センター内

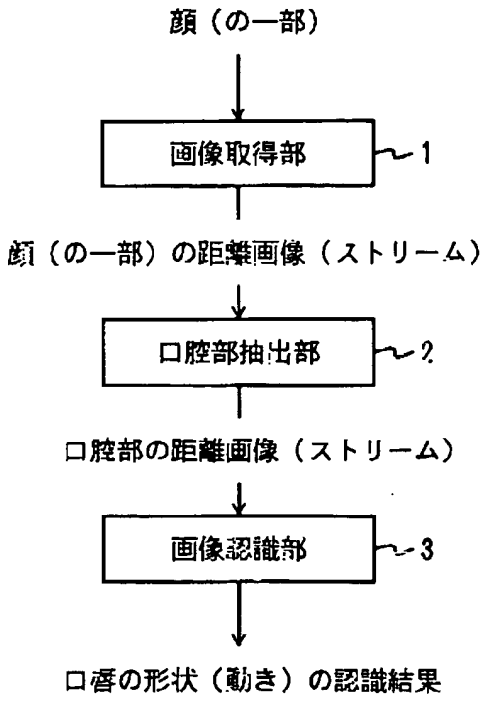
(74)代理人　弁理士　鈴江　武彦　（外６名）

(54)【発明の名称】　画像認識装置及び画像認識装置方法

(57)【要約】

【課題】　人間の口唇の形状や動きを高速かつ高精度に認識可能な画像認識装置を提供すること。

【解決手段】　対象物体に対する距離画像ストリームを取得するための画像取得部と、前記画像取得部により取得された距離画像ストリームから口腔部分を抽出する口腔部抽出部と、前記口腔部抽出部により抽出された口腔部分の距離画像ストリームに基づいて、口唇の形状および口唇の動きの少なくとも一方を認識するための画像認識部とを具備したことを特徴とする。



【特許請求の範囲】
【請求項１】対象物体に対する距離画像を取得するための画像取得手段と、
前記画像取得手段により取得された距離画像から口腔部分を抽出する口腔部抽出手段と、
前記口腔部抽出手段により抽出された口腔部分の距離画像に基づいて、口唇の形状を認識するための画像認識手段とを具備したことを特徴とする画像認識装置。
【請求項２】対象物体に対する距離画像ストリームを取得するための画像取得手段と、
前記画像取得手段により取得された距離画像ストリームから口腔部分を抽出する口腔部抽出手段と、
前記口腔部抽出手段により抽出された口腔部分の距離画像ストリームに基づいて、口唇の形状および口唇の動きの少なくとも一方を認識するための画像認識手段とを具備したことを特徴とする画像認識装置。
【請求項３】対象物体に対する距離画像を取得するための画像取得手段と、
前記画像取得手段により取得された距離画像から顔部分を抽出する顔部抽出手段と、
前記顔部抽出手段により抽出された顔部分の距離画像に基づいて、顔の形状を認識するための画像認識手段とを具備したことを特徴とする画像認識装置。
【請求項４】対象物体に対する距離画像ストリームを取得するための画像取得手段と、
前記画像取得手段により取得された距離画像ストリームから顔部分を抽出する顔部抽出手段と、
前記顔部抽出手段により抽出された顔部分の距離画像ストリームに基づいて、顔の形状および顔の動きの少なくとも一方を認識するための画像認識手段とを具備したことを特徴とする画像認識装置。
【請求項５】前記画像認識手段により得られた前記形状の情報または前記動きの情報に基づいて、話者の顔の向きを識別するための方向識別手段をさらに具備したことを特徴とする請求項１ないし４のいずれか１項に記載の画像認識装置。
【請求項６】前記画像取得手段により取得された距離画像から顔部分を抽出する顔部抽出手段と、
前記顔部抽出手段により抽出された顔部分の距離画像に基づいて、話者の顔の向きを識別するための方向識別手段とをさらに具備したことを特徴とする請求項１または２に記載の画像認識装置。
【請求項７】入力された音声を認識するための音声認識手段と、
前記画像認識手段による認識結果に基づいて話者の会話の開始が検出された場合に前記音声認識手段による音声認識を開始させる制御と前記画像認識手段による認識結果に基づいて話者の会話の終了が検出された場合に前記音声認識手段による音声認識を終了させる制御の少なくとも一方の制御を行う制御手段とをさらに具備したことを特徴とする請求項１、２または６に記載の画像認識装置。
【請求項８】入力された音声を認識するための音声認識手段と、
前記方向識別手段による識別結果が正面である場合に前記音声認識手段による音声認識を開始させる制御と前記方向識別手段による識別結果が正面でない場合に前記音声認識手段による音声認識を終了させる制御の少なくとも一方の制御を行う制御手段とをさらに具備したことを特徴とする請求項５または６に記載の画像認識装置。
【請求項９】所定の出力形態により所定の情報を呈示するための情報呈示手段と、
前記画像認識手段による認識結果に基づいて話者の会話の開始と終了の少なくとも一方の検出を行い、該検出結果に応じて、前記情報呈示手段による情報呈示を開始させる制御と前記情報呈示手段による情報呈示を終了させる制御と前記情報呈示手段により行われている情報呈示に用られている出力形態の少なくとも一部の変更を行う制御のうち少なくとも１つの制御を行う制御手段とをさらに具備したことを特徴とする請求項１、２、６または７に記載の画像認識装置。
【請求項１０】所定の出力形態により所定の情報を呈示するための情報呈示手段と、
前記方向識別手段による識別結果に係る向きと正面方向との関係に応じて、前記情報呈示手段による情報呈示を開始させる制御と前記情報呈示手段による情報呈示を終了させる制御と前記情報呈示手段により行われている情報呈示に用られている出力形態の少なくとも一部の変更を行う制御のうち少なくとも１つの制御を行う制御手段とをさらに具備したことを特徴とする請求項５、６または８に記載の画像認識装置。
【請求項１１】得られた所定の情報を通信するための通信手段をさらに具備したことを特徴とする請求項１ないし１０のいずれか１項に記載の画像認識装置。
【請求項１２】与えられた、対象物体に対する距離画像から、口腔部分を抽出し、
抽出された口腔部分の距離画像に基づいて、口唇の形状を認識することを特徴とする画像認識方法。
【請求項１３】与えられた、対象物体に対する距離画像ストリームから、口腔部分を抽出し、
抽出された口腔部分の距離画像ストリームに基づいて、口唇の形状および口唇の動きの少なくとも一方を認識することを特徴とする画像認識方法。
【請求項１４】コンピュータに、与えられた対象物体に対する距離画像から口腔部分を抽出させ、抽出された口腔部分の距離画像に基づいて口唇の形状を認識させるための手順を含むプログラムを記録したコンピュータ読取り可能な記録媒体。
【請求項１５】コンピュータに、与えられた対象物体に対する距離画像ストリームから口腔部分を抽出させ、抽

出された口腔部分の距離画像ストリームに基づいて口唇の形状および口唇の動きの少なくとも一方を認識させるための手順を含むプログラムを記録したコンピュータ読取り可能な記録媒体。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、取得した距離画像に基づいて画像の形状および／または動きを認識する画像認識装置及び画像認識方法に関する。

【０００２】

【従来の技術】従来、人間の口唇の形状や動きを認識して読唇したり、顔の向き、表情などを判別するような画像処理を行う場合、まず、ＣＣＤカメラなどの撮像装置を用いて、人間の口唇周辺や顔部などを撮影し、その画像から背景などの余計な部分を取り除き、口唇部のみ、顔のみなど認識したい対象のみを切り出すという前処理を行う。そして、その処理後の画像を用いることで、形状や動きなどの認識を行っていた。

【０００３】まず、この認識対象の切り出しという前処理部分について説明する。

【０００４】従来の手法では、カメラで撮影した画像から取得したい対象物の部分のみを切り出す処理において、対象物とそれ以外の部分との間の何らかの相違点を手掛かりとして対象物の切り出しが行われていた。この手掛かりとして、色相の変化を利用する方法、差分画像を利用する方法、マーカーなどを利用する方法、クロマキーを利用する方法などが用いられていた。これらについて、人物の映っている画像から、口唇部分のみを切り出す場合を例として説明する。

【０００５】色相の変化を利用する方法では、口唇の部分はほぼ均一に赤色をしており、周りの肌の部分はほぼ均一に肌色をしている、という色相（画素値）の急激な変化を利用することで、口唇部のみを判別し、切り出しを行っていた。

【０００６】しかしこの方法では、照明の状況によって、肌や口唇の部分に影ができるなどして、色相が変化してしまうなど、通常と異なる色相を示す環境下では、巧く、確実に抽出することが出来なくなるといったような問題点があった。また、口唇の形状を安定的に得るために、特定の色の口紅を用いることで、色相変化を強調したりしなければならない場合もあった。

【０００７】差分画像を利用する方法では、話者が会話をしている際には、顔の中で、口唇の部分のみが動いているということを利用して、現在のフレームと、次のフレームとの差分画像を取ることによって、動いている部分を取得し、それを口唇の部分とする、ということが行われていた。

【０００８】しかしこの方法では、背景で何かが動いているような環境下では、口唇以外の不必要な部分も抽出してしまう、口唇が動いていないときには抽出できない、というように、環境や条件に著しく依存してしまい、常に、確実に口唇の部分のみを抽出するのは大変困難であった。

【０００９】マーカーを利用する方法では、口唇の周りに幾つかのマーカーを貼って特徴点とし、その特徴点の動きをもとに、口唇部を抽出していた。

【００１０】しかしこの方法では、顔に、マーカーなどを貼らなくてはならないため、使える環境が限られているなどの問題があった。

【００１１】クロマキーを利用する方法では、例えば、青色など、人物の顔にあまり現れないような色のスクリーンの前に人物が配置し、カメラなどで得た画像から青色を取り除くことで、顔の部分のみを抽出していた。

【００１２】しかしこの方法では、背景の色を強要されるため、特定の状況でのみしか用いることができない、口唇のような顔の内部の一部分のみの抽出ができない、などというような問題があった。

【００１３】このように従来の手法では、カメラで撮影した画像から取得したい対象物の部分のみを確実に切り出す処理は、大変困難なものであった。

【００１４】次に、対象物が切り出された画像から、対象物の形状、動きなどの認識を行う部分について説明する。

【００１５】従来、切り出された対象物の画像は、２次元情報しか含んでいない。これは、従来の撮像装置では３次元形状を取得することは困難であり、３次元形状を取得するような撮像装置があっても、それらは、動きの様なリアルタイムの認識に適していなかったからである。また、そのような３次元形状の撮像装置は、大変高価で、気軽に用いることができないという問題もあった。そのため、従来の画像処理では、２次元情報のみを用いて、人間の顔や口唇の形状、動きといった、本来は３次元的ものを、なんとか認識しようと努力していた。

【００１６】しかし、本来３次元的な形状や動きを２次元情報として用いていたため、必要な情報が欠落してしまい、様々な工夫はしているものの、簡単な形状や動きの認識のみしか行えないといったように、どうしても無理があった。

【００１７】また、上述したとおり、画像から対象物のみを切り出すという作業を確実に行うことは大変困難であるため、この切り出しの不確実さも、認識率を下げる要因に大きく関わっていた。

【００１８】以上のように、従来方法では、画像からの対象物の抽出方法にも、画像の認識方法にも、様々な問題点があった。

【００１９】

【発明が解決しようとする課題】以上のように、従来、カメラで撮影した画像から取得したい対象物の部分のみを確実に切り出す処理は大変困難なものであり、それが画像認識の認識率の低下の要因となっていた。

【００２０】また、様々な制約から、カメラなどを用いて画像を２次元情報として取得していたため、３次元形状や３次元的動きの認識を２次元情報のみから行うしかなく、簡単な形状、動きの認識しか行うことができないという問題があった。

【００２１】本発明は、上記事情を考慮してなされたものであり、人間の顔や口唇の形状や動きを高速かつ高精度に認識可能な画像認識装置及び画像認識装置方法を提供することを目的とする。

【００２２】

【課題を解決するための手段】本発明１（請求項１）に係る画像認識装置は、対象物体に対する距離画像を取得するための画像取得手段と、前記画像取得手段により取得された距離画像から口腔部分を抽出する口腔部抽出手段と、前記口腔部抽出手段により抽出された口腔部分の距離画像に基づいて、口唇の形状を認識するための画像認識手段とを具備したことを特徴とする。

【００２３】本発明２（請求項２）に係る画像認識装置は、対象物体に対する距離画像ストリーム（距離画像の動画像）を取得するための画像取得手段と、前記画像取得手段により取得された距離画像ストリーム（距離画像の動画像）から口腔部分を抽出する口腔部抽出手段と、前記口腔部抽出手段により抽出された口腔部分の距離画像ストリームに基づいて、口唇の形状および口唇の動きの少なくとも一方を認識するための画像認識手段とを具備したことを特徴とする。

【００２４】本発明によれば、対象物体に対する距離画像から必要とする部分を抽出し、抽出した部分の距離画像に基づいて認識処理を行うので、話者の口唇認識（例えば、口腔部の形状、動きや、発言内容の認識など）等を高速かつ高精度に行うことができる。

【００２５】本発明３（請求項３）に係る画像認識装置は、対象物体に対する距離画像を取得するための画像取得手段と、前記画像取得手段により取得された距離画像から顔部分を抽出する顔部抽出手段と、前記顔部抽出手段により抽出された顔部分の距離画像に基づいて、顔の形状を認識するための画像認識手段とを具備したことを特徴とする。

【００２６】本発明４（請求項４）に係る画像認識装置は、対象物体に対する距離画像ストリーム（距離画像の動画像）を取得するための画像取得手段と、前記画像取得手段により取得された距離画像ストリーム（距離画像の動画像）から顔部分を抽出する顔部抽出手段と、前記顔部抽出手段により抽出された顔部分の距離画像ストリームに基づいて、顔の形状および顔の動きの少なくとも一方を認識するための画像認識手段とを具備したことを特徴とする。

【００２７】本発明によれば、対象物体に対する距離画像から必要とする部分を抽出し、抽出した部分の距離画像に基づいて認識処理を行うので、話者の顔部認識（例えば、顔部の形状、動きの認識など）等を高速かつ高精度に行うことができる。

【００２８】本発明５（請求項５）は、請求項１ないし４のいずれか１項に係る画像認識装置において、前記画像認識手段により得られた前記形状の情報または前記動きの情報に基づいて、話者の顔の向きを識別するための方向識別手段をさらに具備したことを特徴とする。

【００２９】本発明６は、請求項１または２に係る画像認識装置において、前記画像認識手段により認識された口唇の形状もしくは口唇の動きに基づいて、話者の顔の向きを識別するための方向識別手段をさらに具備したことを特徴とする。

【００３０】本発明７は、請求項３または４に係る画像認識装置において、前記画像取得手段により取得された顔の形状もしくは顔の動きに基づいて、話者の顔の向きを識別するための方向識別手段をさらに具備したことを特徴とする。

【００３１】本発明８（請求項６）は、請求項１または２に係る画像認識装置において、前記画像取得手段により取得された距離画像から顔部分を抽出する顔部抽出手段と、前記顔部抽出手段により抽出された顔部分の距離画像に基づいて、話者の顔の向きを識別するための方向識別手段とをさらに具備したことを特徴とする。

【００３２】本発明によれば、話者の口唇認識あるいは顔部認識等と伴に、話者の向いている方向の識別をすることができる。

【００３３】本発明９（請求項７）は、請求項１、２または６に係る画像認識装置において、入力された音声を認識するための音声認識手段と、前記画像認識手段による認識結果に基づいて話者の会話の開始が検出された場合に前記音声認識手段による音声認識を開始させる制御と前記画像認識手段による認識結果に基づいて話者の会話の終了が検出された場合に前記音声認識手段による音声認識を終了させる制御の少なくとも一方の制御を行う制御手段とをさらに具備したことを特徴とする。

【００３４】本発明１０（請求項８）は、請求項５または６に係る画像認識装置において、入力された音声を認識するための音声認識手段と、前記方向識別手段による識別結果が正面である場合に前記音声認識手段による音声認識を開始させる制御と前記方向識別手段による識別結果が正面でない場合に前記音声認識手段による音声認識を終了させる制御の少なくとも一方の制御を行う制御手段とをさらに具備したことを特徴とする。

【００３５】本発明によれば、話者の口唇認識あるいは顔部認識等と伴に、口唇認識結果あるいは話者の向いている方向に応じた音声認識の制御を行うことができる。

【００３６】本発明１１（請求項９）は、請求項１、２、６または７に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示す

るための情報呈示手段と、前記画像認識手段による認識結果に基づいて話者の会話の開始と終了の少なくとも一方の検出を行い、該検出結果に応じて、前記情報呈示手段による情報呈示を開始させる制御と前記情報呈示手段による情報呈示を終了させる制御と前記情報呈示手段により行われている情報呈示に用られている出力形態の少なくとも一部の変更を行う制御のうち少なくとも１つの制御を行う制御手段とをさらに具備したことを特徴とする。

【００３７】本発明１２（請求項１０）は、請求項５、６または８に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記方向識別手段による識別結果に係る向きと正面方向との関係に応じて、前記情報呈示手段による情報呈示を開始させる制御と前記情報呈示手段による情報呈示を終了させる制御と前記情報呈示手段により行われている情報呈示に用られている出力形態の少なくとも一部の変更を行う制御のうち少なくとも１つの制御を行う制御手段とをさらに具備したことを特徴とする。

【００３８】本発明によれば、話者の口唇認識あるいは顔部認識等と伴に、口唇認識結果や話者の向いている方向に応じた情報呈示の制御を行うことができる。

【００３９】本発明１３は、請求項１、請求項２、発明６または請求項６に係る画像認識装置において、入力された音声を認識するための音声認識手段と、前記画像認識手段による認識結果に基づいて話者の会話（話者の行為実施）の開始を検出し、該会話の開始が検出された場合に前記音声認識手段による音声認識を開始させる音声認識開始手段とをさらに具備したことを特徴とする。

【００４０】本発明１４は、請求項１、請求項２、発明６、請求項６または発明１３に係る画像認識装置において、入力された音声を認識するための音声認識手段と、前記画像認識手段による認識結果に基づいて話者の会話（話者の行為実施）の終了を検出し、該会話の終了が検出された場合に前記音声認識手段による音声認識を終了させる音声認識終了手段とをさらに具備したことを特徴とする。

【００４１】本発明１５は、請求項５、発明６、発明７または請求項６に係る画像認識装置において、入力された音声を認識するための音声認識手段と、前記方向識別手段による識別結果が正面（話者の行為実施）である場合に、前記音声認識手段による音声認識を開始させる音声認識開始手段とをさらに具備したことを特徴とする。

【００４２】本発明１６は、請求項５、発明６、発明７、請求項６または発明１５に係る画像認識装置において、入力された音声を認識するための音声認識手段と、前記方向識別手段による識別結果が正面（話者の行為実施）でない場合に、前記音声認識手段による音声認識を終了させる音声認識終了手段とをさらに具備したことを特徴とする。

【００４３】本発明１７は、請求項１、請求項２、発明６、請求項６、発明１３または発明１４に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記画像認識手段による認識結果に基づいて話者の会話（話者の行為実施）の開始を検出し、該会話の開始が検出された場合に前記情報呈示手段による情報呈示を開始させる情報呈示開始手段とをさらに具備したことを特徴とする。

【００４４】本発明１８は、請求項１、請求項２、発明６、請求項６、発明１３、発明１４または発明１８に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記画像認識手段による認識結果に基づいて話者の会話（話者の行為実施）の終了を検出し、該終了が検出された場合に前記情報呈示手段による情報呈示を終了させる情報呈示終了手段とをさらに具備したことを特徴とする。

【００４５】本発明１９は、請求項５、発明６、発明７、請求項６、発明１５または発明１６に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記方向識別手段による識別結果が正面（話者の行為実施）である場合に、前記情報呈示手段による情報呈示を開始させる情報呈示開始手段とをさらに具備したことを特徴とする。

【００４６】本発明２０は、請求項５、発明６、発明７、請求項６、発明１５、発明１６または発明２０に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記方向識別手段による識別結果が正面（話者の行為実施）でない場合に、前記情報呈示手段による情報呈示を終了させる情報呈示終了手段とをさらに具備したことを特徴とする。

【００４７】本発明２１は、請求項１、請求項２、発明６、請求項６、発明１３、発明１４、発明１７または発明１８に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記画像認識手段による認識結果に基づいて話者の会話（話者の行為実施）の開始を検出し、該会話の開始が検出された場合に、前記情報呈示手段による情報呈示をそれまでとは異なる出力形態による情報呈示に切り替える情報呈示切り替え手段とをさらに

具備したことを特徴とする。

【００４８】本発明２２は、請求項１、請求項２、発明６、請求項６、発明１３、発明１４、発明１７、発明１８または発明２１に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記画像認識手段による認識結果に基づいて話者の会話（話者の行為実施）の終了を検出し、該会話の終了が検出された場合に、前記情報呈示手段による情報呈示をそれまでとは異なる出力形態による情報呈示に切り替える情報呈示切り替え手段とをさらに具備したことを特徴とする。

【００４９】本発明２３は、請求項５、発明６、発明７、請求項６、発明１５、発明１６、発明１９または発明２０に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記方向識別手段による識別結果が正面（話者の行為実施）となった場合に、前記情報呈示手段による情報呈示をそれまでとは異なる出力形態による情報呈示に切り替える情報呈示切り替え手段とをさらに具備したことを特徴とする。

【００５０】本発明２４は、請求項５、発明６、発明７、請求項６、発明１５、発明１６、発明１９、発明２０または発明２３に係る画像認識装置において、所定の出力形態（音声、画像、あるいは他の形態、あるいは複数の形態を組み合わせたもの）により所定の情報を呈示するための情報呈示手段と、前記情報呈示手段による情報呈示中に前記方向識別手段による識別結果が正面（話者の行為実施）でなくなった場合に、前記情報呈示手段による情報呈示をそれまでとは異なる出力形態による情報呈示のみに切り替える情報呈示切り替え手段とをさらに具備したことを特徴とする。

【００５１】本発明２５（請求項１１）は、請求項１ないし１０のいずれか１項に係る画像認識装置において、得られた所定の情報を通信するための通信手段をさらに具備したことを特徴とする。

【００５２】本発明によれば、認識結果等の所望の情報を外部の装置に与えることができる。

【００５３】本発明２６は、請求項１または２に係る画像認識装置において、前記画像認識手段により認識された口唇の形状の情報もしくは口唇の動きの情報を通信するための通信手段をさらに具備したことを特徴とする。

【００５４】本発明２７は、請求項３または４に係る画像認識装置において、前記画像認識手段により認識された顔の形状の情報もしくは顔の動きの情報を通信するための通信手段をさらに具備したことを特徴とする。

【００５５】本発明２８は、請求項５、発明６、発明７または請求項６に係る画像認識装置において、前記方向識別手段により識別された話者の顔の向きの情報を通信するための通信手段をさらに具備したことを特徴とする。

【００５６】本発明２９は、請求項７、請求項８、発明８、発明９、発明１０または発明１１に係る画像認識装置において、前記音声認識手段による認識結果を通信するための通信手段をさらに具備したことを特徴とする。

【００５７】本発明３０は、請求項９、請求項１０、発明１２、発明１３、発明１４、発明１５、発明１６、発明１７、発明１８または発明１９に係る画像認識装置において、前記情報呈示手段により呈示された情報を通信するための通信手段をさらに具備したことを特徴とする。

【００５８】なお、以上の各本発明において、画像取得手段を省き、対象物体に対する距離画像を外部から与えるようにした構成も成立する。

【００５９】本発明３１（請求項１２）に係る画像認識方法は、与えられた、対象物体に対する距離画像から、口腔部分を抽出し、抽出された口腔部分の距離画像に基づいて、口唇の形状を認識することを特徴とする。

【００６０】本発明３２（請求項１３）に係る画像認識方法は、与えられた、対象物体に対する距離画像ストリームから、口腔部分を抽出し、抽出された口腔部分の距離画像ストリームに基づいて、口唇の形状および口唇の動きの少なくとも一方を認識することを特徴とする。

【００６１】本発明３３（請求項１４）は、コンピュータに、与えられた対象物体に対する距離画像から口腔部分を抽出させ、抽出された口腔部分の距離画像に基づいて口唇の形状を認識させるための手順を含むプログラムを記録したコンピュータ読取り可能な記録媒体を要旨とする。

【００６２】本発明３４（請求項１５）は、コンピュータに、与えられた対象物体に対する距離画像ストリームから口腔部分を抽出させ、抽出された口腔部分の距離画像ストリームに基づいて口唇の形状および口唇の動きの少なくとも一方を認識させるための手順を含むプログラムを記録したコンピュータ読取り可能な記録媒体を要旨とする。

【００６３】なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

【００６４】また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても成立する。

【００６５】

【発明の実施の形態】以下、図面を参照しながら発明の実施の形態を説明する。

【００６６】（第１の実施形態）まず、本発明の第１の実施形態について説明する。
【００６７】図１は、本発明の第１の実施形態に係る画像認識装置の全体構成図である。
【００６８】本実施形態の画像認識装置は、距離画像ストリームを取得するための画像取得部１と、画像取得部１で取得された顔の全部または一部の距離画像ストリームから、口腔部分のみを抽出する口腔部抽出部２と、抽出された口腔部の距離画像ストリームから、口唇の形状および／または口唇の動きを認識する画像認識部３とから構成される。
【００６９】画像取得部１は、画像認識対象物体となる人間の顔の全部または一部を、その３次元形状を反映した奥行き値を持つ画像（以下、距離画像と呼ぶ）として所定時間毎（例えば１／６０秒毎など）に取得するものである（例えば特願平９－２９９６４８の画像取得方法を用いて実現することができる）。画像取得部１は概略的には、例えば、対象物体に光を照射し、対象物体からの反射光の空間的な強度分布を抽出し、その各画素の強度値を奥行きあるいは距離を示す値に変換することにより、距離画像を生成する。この画像取得部１を用いて顔を撮像することで、顔の全部または一部分の、距離画像による動画像（以下、距離画像ストリームと呼ぶ）を得ることができる。なお、画像取得部１の詳細については後述する。
【００７０】図２に、画像取得部１により取得された顔の距離画像（距離画像ストリーム中の１フレーム分）の例を示す。距離画像は、奥行き情報を有する３次元画像で、例えば、ｘ軸（横）方向６４画素、ｙ軸（縦）方向６４画素、ｚ軸（奥行き）方向２５６階調の画像になっている。図２は、距離画像の距離値すなわちｚ軸方向の階調をグレースケールで表現したものである。距離画像においては、色が白に近いほど距離が近く、黒に近くなるほど距離が遠い。また、色が完全に黒のところは、画像がない、あるいはあっても遠方でないのと同じであることを示している。例えば、図２は、口唇部が白く、その内側の口腔部が黒くなっている様子を示すものである。
【００７１】なお、画像取得部１における受光面もしくはこれを収容した筐体は、本画像認識装置の目的等に応じて適宜設置するばよい。例えば本画像認識装置が表示装置を持つものである場合、この表示装置に対して対象物体となる人間の顔が正面を向いたときに、当該受光面に対しても正面を向いた形になるように当該画像認識装置の筐体に設ける。
【００７２】次に、口腔部抽出部２について説明する。
【００７３】口腔部抽出部２は、画像取得部１によって取得された顔の全部または一部の距離画像ストリームから、口腔部のみを抽出するものである。
【００７４】人間の口唇の周辺部分を３次元的に見た場合、その局所的な形状は人によって様々であるし、同じ人でも状況によって様々な形状をしている。しかし、大局的には、「口唇部が少し凸形状をしており、その内側の口腔部が大きく凹形状をしている」という、人や状況に依らず一意に定まる特徴がある。
【００７５】図３は口唇を閉じている場合の顔の距離画像を、図４は口唇を開いている場合の顔の距離画像を、それぞれ、立体的に示したものである。図３および図４を見ると、上述したような口腔部の３次元的特徴がはっきりと見て取れることが分かる。
【００７６】この口腔部の３次元形状の特徴を積極的に利用すれば、顔の距離画像ストリームから、口腔部のみを抽出した距離画像ストリームを構成することは容易である。
【００７７】以下では、口腔部抽出部２でどのように口腔部を抽出するのかを具体的に説明する。
【００７８】画像取得部１によって取得された距離画像（以下、原画像とも呼ぶ）は、顔の３次元的形状を表している。この距離画像の２階微分画像を求めることで、原画像における傾き変化の様子を知ることができる。これを用いれば、原画像のエッジ部分を抽出することができる。なお、ここでエッジと言うのは、顔と背景との境界や、口唇と肌との境界のように、傾きの変化がある部分のことである。
【００７９】図５にエッジ抽出の具体的な処理の流れの一例を示す。
【００８０】まず、Ｍａｒｒ－Ｈｉｌｄｒｅｔｈが提案したガウスラプラシアンフィルタを原画像に施す（ステップＳ１００）。
【００８１】次に、そのゼロクロス点を求める（ステップＳ１０１）。このとき、例えば、注目画素の４近傍の画素値が正である点をゼロクロス点とすればよい。
【００８２】そして、ゼロクロス点ならば、図６に示すようなＳｏｂｅｌオペレータ（図中（ａ）がＸ方向に対応し、（ｂ）がＹ方向に対応する）を施し、その画素の強度を求める（ステップＳ１０２）。
【００８３】この強度がある閾値以上ならば、エッジの構成点であるとみなす（ステップＳ１０３）。
【００８４】以上の処理により、原画像から、エッジ部分のみを抽出することができる。
【００８５】なお、ここでは、エッジ抽出の一手法として、ガウスラプラシアンフィルタ、Ｓｏｂｅｌオペレータを用いる方法について説明したが、これに限定されるものではなく、ハフ変換を用いる方法など、別の手法を用いて実現しても良い。
【００８６】以上で説明した処理を距離画像に施すことで、顔の距離画像から、エッジ部分のみを抽出することができる。さらに、このエッジ情報と、口唇の形状（ループ状（穴）のエッジを持つもののなかで、一番大きなものなど）の情報を用いることで、口唇部のエッジのみ

を抽出することができる。

【0087】この方法では、実際の顔の3次元形状をもとに、エッジの抽出を行っているため、従来の2次元画像から色相の変化などを利用してエッジを抽出する方法と比べて、エッジの誤認識（余分なエッジの抽出）をすることがなく、確実に口腔部のみを切り出すことが可能である。これは、3次元形状は実際のエッジに深く関係しているのに対し、色相変化を用いる方法は色相が異なる部分をエッジと見なして判断する一手段ではあるが、決定的なものではないからである。

【0088】以上の処理で、顔の距離画像ストリームから、口唇部のみの距離画像ストリームを取得することができる。

【0089】なお、ここでは、顔の距離画像から、口腔部を抽出する方法として、傾きの変化を利用する方法について説明したが、これに限定されるものではない。例えば、口腔部の「窪み」という幾何学的な形状（奥行きZ値が一定値以下）を利用して、閾値を設けることで「窪み」部分を抽出してもよいし、幾何学的推論を行うことによって抽出しても良い。また、口腔部の「窪み」状のテンプレートをあらかじめ用意しておいて、それとのパターンマッチングを取ることで求めてもよい。また、距離情報を用いてバンドパスフィルタによるフィルタリング処理を行うことでもエッジを取ることができる。他の3次元形状を利用して抽出する方法でも構わない。

【0090】次に、画像認識部3について説明する。

【0091】画像認識部3は、口腔部抽出部2によって抽出された口腔部の距離画像ストリームをもとに、口唇の形状および／または動きを認識するものである。

【0092】まず、口唇の形状の認識について説明する。

【0093】画像認識部3では、「あ」、「い」、…、といった様々なテンプレートを予め用意しておき、それらと口腔部抽出部2で得られた口唇の形状とを比較して、類似度を計算し、類似度の最も高いものを認識結果として採用するという、テンプレートマッチングなどを用いて、認識を行う。

【0094】図7に画像認識部3におけるテンプレートマッチングの処理の流れの一例を示す。

【0095】まず、抽出された口腔部の距離画像（原画像）を、テンプレートの方向、サイズに合わせて正規化する（ステップS200）。

【0096】次に、用意した様々なテンプレートの中から、原画像と比較すべきテンプレートkを選択する（ステップS201）。

【0097】次に、原画像とテンプレートとのハミング距離を計算する（ステップS202）。ハミング距離（H）は、例えば、$H=\Sigma_i \Sigma_j | d(i,j) - t_k(i,j) |$により計算する。ここで、i、jはそれぞれ各画素のx、y座標、d（i，j）は原画像の座標（i，j）での距離値、t_k（i，j）はテンプレートkの座標（i，j）での距離値である。

【0098】なお、ここでは、ハミング距離の導出の一方法を説明したが、ハミング距離の導出は、これに限定されるものではなく、他の計算式を用いても良い。

【0099】これらの処理を全てのテンプレートについて行うため、全てのテンプレートについて、上述のハミング距離の計算が終了しているか判定する（ステップS203）。

【0100】未だハミング距離の計算が終わっていないテンプレートがあれば、ステップS201に戻る。

【0101】全てのテンプレートについて、原画像とのハミング距離の計算が終了したら、それらを比較し、最も値の小さなテンプレートを見つける。そして、このテンプレートの表現している内容を認識結果とする（ステップS204）。例えば、この選ばれたテンプレートが、「た」を発音している際の口唇形状であったならば、原画像の距離画像の発音（口唇形状）は「た」であったと認識する。

【0102】以上の処理を距離画像ストリームに含まれる、全ての距離画像に対して、順次行うことによって、話者の発話内容の認識が行われる。

【0103】なお、以下では、音声認識と区別するために、口唇形状から話者の発話内容を認識すること（認識対象となった者が現実には音声を出さず、実際に話すときと同じように口唇を動した場合に得られた距離画像に基づく認識を含む）を口唇認識と呼ぶ。

【0104】次に、口唇の動きの認識について説明する。

【0105】口唇の動きの認識を行う場合、例えば、「口を開け閉めしている」、「あくびをしている」といったような、動きを表すテンプレートの列（動きを各フレームに分割し、それぞれを1つのテンプレートとして、一連の動きのテンプレートをまとめたもの）を用意しておき、上述したものと同様に、距離画像ストリームに含まれる全ての距離画像に対して、前記テンプレートの列と順次テンプレートマッチングを行うことで、動きに対する口唇認識を行うこともできる。

【0106】以上のような方法で得られた口唇認識の結果は、従来の画像認識と異なり、実際の口唇の3次元形状を利用することによって、認識を行った結果である。従来は、通常のビデオカメラの画像などから抽出した2次元的な口唇形状を用いて認識していたため、口唇の平面的な動きのみから認識を行うしかなかったが、この方法では、上述の通り、3次元の情報を用いることが可能であるため、従来よりも、より多くの情報を用いて認識することが可能である。そこで、正面から見たときの口唇形状がほぼ同じで、口唇の奥行き方向の形状が異なっているというような、従来なら認識が不可能であった場

合も、本実施形態の画像認識装置を用いることで認識することが可能となっている。また、識別する手掛かりが増えているため、従来よりも、認識率も高くなり、誤認識し難いという利点もある。

【0107】なお、ここでは、原画像とテンプレートとのハミング距離を求めることで、原画像とテンプレートの類似度を計算する方法について説明したが、類似度の計算は、これに限定されるものではない。DPマッチング法、KL変換法などを用いて求める方法、原画像をフーリエ変換し、フーリエ変換後の画像について相関関係を求めることで、類似度を計算する方法など、あらゆる方法を用いることができる。

【0108】また、ここでは、口腔部の距離画像ストリームから、口唇の形状、動きを認識する方法として、テンプレートマッチングを行う方法について説明したが、これに限定されるものではなく、例えば、口唇の形状から、筋肉の動きを求めて、その形状変化を手掛かりとして、筋肉モデルから発音内容を類推する、などのように他の方法で認識を行ってもよい。

【0109】以上のように本実施形態によれば、口唇の距離画像を用いることで、あまり計算コストをかけずに、容易に、口唇部を抽出することが可能となる。さらに、口唇認識に関しても、抽出した口唇部の3次元形状の情報を用いることにより、従来方法では、判別に難しかった（誤認識が多かった）ような形状に関する認識や、従来では不可能であったような形状に関する認識が可能になる。

【0110】以上のようにして得た口唇の形状の認識結果、口唇の動きの認識結果、あるいは口唇の形状の認識結果と口唇の動きの認識結果を組み合わせたものは、その後の種々の処理に供することができる。なお、画像認識部3に、口唇の形状と動きの認識の両方の機能を設けるか、いずれか一方を設けるかは、システムの目的等に応じて適宜設計することが可能である。

【0111】本実施形態は、上記した構成に限定されず、種々変形して実施することができる。以下では、本実施形態のいくつかの変形例を示す。

【0112】（第1の実施形態の変形例1）口唇部抽出部2の代わりに、画像取得部1で所得された距離画像ストリームから顔部分のみを抽出するための顔部抽出部を具備してもよい。

【0113】そして、画像認識部3で、予めA氏、B氏、というように人物の顔形状のテンプレートを用意しておき、それらを用いて顔部抽出部5で抽出された顔部の距離画像とのマッチングを行うことで、本実施形態の画像認識装置で撮像された人物が誰であるのかを認識することができる。

【0114】これにより、例えば、本実施形態の画像認識装置（または少なくとも画像取得部1の発光素子と受光素子の部分）を、自動ドアの近くなどに置き、そこを通る人物の顔を認識することで、特定の人物と認識したときのみドアを開ける、といったような、簡単なセキュリティチェックに使うことが可能である。

【0115】（第1の実施形態の変形例2）本実施形態は、医療面でも重病者の看護に有効である。従来、病室や在宅看護者の家庭などにいる患者が何か異常をきたした場合には、枕元にある押しボタン式のブザーで、看護婦や医者に知らせていた。しかし、押しボタン式のブザーでは、患者が弱っていた場合に、ボタンを押す余裕が無いことが多く、危険であった。このような場所に第1の実施形態の画像認識装置を置くことで、病気で弱っていて、あまり声を出せないような場合でも、病人のわずかな声と、微妙な口唇の動きから、病人が何か伝えたいということを判別することが可能である。

【0116】これを押し進めて、普段口唇を動かすことがない病人が口唇を動かしたら、病状が急変した可能性がある。このような場合には、口唇の動きを何らかの音に変換して、警報音代わりに用いることができ、それにより医者や看護婦が病室や在宅看護者の家庭に駆け付けるような方策をとることができる。

【0117】このような場合、図8に例示するように、口唇認識の結果をそのまま音声に変換し呈示する、または、結果に応じて何らかの音を呈示するを音呈示部4を設ける。

【0118】（第1の実施形態の変形例3）図9に例示するように、上記の第1の実施形態の変形例2の構成（図8）に、さらに顔部のみの距離画像ストリームを抽出するための顔部抽出部5を付加して、顔の3次元形状情報を用いることで、例えば、顔を上下に振っているなどというように、顔のゼスチャーの認識を行ったり、笑っている、怒っている、困っているなどというように、表情の認識を行うことが可能である。

【0119】その際、画像認識部3では、例えば、頷く：顔を上下に数回振る、拒む：顔を左右に数回振る、喜ぶ：大きく口があく、目が細くなる、驚く：目を見開く、などというようにゼスチャーや表情などを得るためのテンプレートを用意しておき、それらを用いてテンプレートマッチングを行うことで、顔のゼスチャーや表情の認識を行う。

【0120】そして、この認識した表情に応じて、口唇の動きを音声変換する際に、変換する音声の種類やピッチなどを変えることも可能である。

【0121】また、例えば、同じ口唇の動きでも、肯定の場合は犬のなき声、否定の場合はニワトリの鳴き声、喜んでいる場合は猫のなき声というように変化させることもできる。このようにすることで、例えば、子供に、英語の単語発生などを楽しく飽きないように勉強できるようにすることが可能となる。

【0122】（第2の実施形態）次に、本発明の第2の実施形態について説明する。本実施形態では、第1の実

施形態と相違する部分を中心に説明する。

【0123】図10は、本発明の第2の実施形態に係る画像認識装置の全体構成図である。

【0124】図10に示されるように、本実施形態の画像認識装置は、第1の実施形態の画像認識装置の構成に対して、画像認識部3で得られた口唇の形状もしくは動きの認識結果をもとに、話者の顔の向いている方向を識別するための方向識別部6が追加された構成になっている。

【0125】これにより話者の発言内容だけでなく、同時に、話者がどちらの方向を向いて話しているかを認識することができる。

【0126】次に、方向識別部6について説明する。

【0127】方向識別部6では、画像認識部3で得られた口唇の形状もしくは動きの認識結果をもとに、話者の顔の向いている方向を識別する。その際、口唇の3次元形状を利用することで、話者の顔の向きを計算する。

【0128】以下では、話者の顔の向いている方向を求める具体的な方法の一例について、図11に示す処理の流れ図を用いて説明する。

【0129】まず、口唇の距離画像中のある画素X（例えば座標値（i、j））を選択する（ステップS300）。

【0130】次に、画素Xと隣接している画素Y（例えば座標値（i－1、j））を選択する（ステップS301）。

【0131】次に、図12（（a）は隣接8画素を示す図、（b）は傾きベクトルgとこれに直交する法線ベクトルpを説明するための図）のように、選択した画素Y（例えば座標値（i－1、j））との距離値の差d（i、j）－d（i－1、j）をもとに、この2画素間の傾きベクトルgを求める（ステップ302）。

【0132】この2画素X、Yと同一平面上にあり、ステップS302で得られた傾きベクトルgと直行する法線ベクトルpを求める（ステップS303）。

【0133】画素Xと隣接する全ての画素Yについて法線ベクトルの計算が終了したか判別する（ステップS304）。

【0134】全ての隣接画素について終了していなかったら、ステップS301に戻る。全てについて終了していたら、この法線ベクトルの平均P＝Σpを計算し、画素Xの法線ベクトルPとする（ステップS305）。

【0135】以上の処理を距離画像中の全ての画素について行ったかどうか判定する（ステップS306）。行っていなかったら、ステップS300に戻る。

【0136】全ての画素について、法線ベクトルPの計算が終了したら、各画素の法線ベクトルの平均P_{lip}＝Σpを計算し、これを口唇の法線ベクトルとする（ステップS307）。

【0137】口唇は、顔のほぼ中央にあり、ほぼ左右上下対称形状であるため、口唇の法線ベクトルと顔の法線ベクトルの方向は、おおむね一致する。そのため、ステップS307で得られたP_{lip}が顔の法線ベクトルとなる。つまり、法線ベクトルP_{lip}を顔の向きとして話者の向いている方向を識別することができる。

【0138】なお、ここでは、口唇の向いている方向を得る一手段として、距離画像から口唇の法線ベクトルを計算する方法について説明したが、これに限定されるものではなく、口唇の大きさの比率や形状の変化から口唇の向いている方向を類推するなど、他の方法を用いても構わない。

【0139】以上のように本実施形態によれば、話者がどちらの方向を向いて、どのような話をしているのかもしくはどのような口唇の動きをしているのかなどを、同時に認識することが可能である。

【0140】本実施形態は、上記した構成に限定されず、種々変形して実施することができる。以下では、本実施形態のいくつかの変形例を示す。

【0141】（第2の実施形態の変形例1）図13のように、口腔部抽出部2の代わりに、画像取得部1で取得された顔の全部または一部の距離画像ストリームから顔部のみを抽出するための顔部抽出部5を置いても良い。この場合、画像認識部3には、顔部抽出部5で抽出された顔部の距離画像ストリームが入力される。

【0142】そして、画像認識部3では、例えば、頷く：顔を上下に数回振る、拒む：顔を左右に数回振る、喜ぶ：大きく口があく、目が細くなる、驚く：目を見開く、などというようにゼスチャーや表情などを得るためのテンプレートを用意しておき、それらを用いて、入力された顔部の距離画像ストリームとのテンプレートマッチングを行うことで、頷いているなどのゼスチャーや、喜んでいる、驚いている、困っているなどの表情変化などを認識することが可能である。

【0143】方向識別部6では、画像認識部3で得られた顔部の形状、動きの認識結果をもとに、話者の顔の向いている方向を識別する。

【0144】このように変形することにより、対象人物が、どちらの方向を向いて、どのような顔の動作（ゼスチャー、表情変化など）をしているのかを認識することができる。

【0145】（第2の実施形態の変形例2）なお、第2の実施形態では、画像認識部3の認識結果をもとに、前記方向識別部6で話者の向いている方向を識別したが、図14のように、画像取得部1で取得された顔の距離画像ストリーム（これには、背景などが含まれる）から顔の部分のみを抽出するための顔部抽出部5を新たに追加し、顔部抽出部5で抽出された顔の距離画像ストリームをもとに、方向識別部6で話者の向いている方向を識別するようにしても良い。この場合、方向識別部6では、顔部抽出部5で抽出された顔の距離画像ストリームか

ら、顔の法線方向（例えば、顔を構成する画素の法線方向の平均）を計算することにより、話者の向いている方向を得る。

【0146】このようにすることで、第2の実施形態では、口唇の向いている方向から、顔の向いている方向を得ていたが、直接、顔の向いている方向を得ることができるため、より細かく、微妙な顔の向きを得ることが可能である。

【0147】（第3の実施形態）次に、本発明の第3の実施形態について説明する。本実施形態では、第2の実施形態と相違する部分を中心に説明する。

【0148】図15は、本発明の第3の実施形態に係る画像認識装置の全体構成図である。

【0149】図15に示されるように、本実施形態の画像認識装置は、第2の実施形態の画像認識装置の構成もしくはその変形例の構成に対して、話者の発言内容を認識する音声認識部7と、方向識別部6で得られた話者の顔の向いている方向をもとに、音声認識部7に、音声認識の開始を指示するための音声認識開始部8が追加された構成になっている。

【0150】これにより話者の顔の向いている方向に応じて、音声認識を行うことができる。

【0151】次に、音声認識部7について説明する。

【0152】音声認識部7は、マイクなどの音声の入力装置を用いて入力された音声の内容を認識するものである。音声認識部7では、種々の認識手法を用いることが可能である。例えば、隠れマルコフモデルなどを用いて実現してもよい。音声認識を行うことで、話者の会話の内容を認識することができる。

【0153】次に、音声認識開始部8について説明する。

【0154】音声認識開始部8は、方向識別部6で得られた結果をもとに、音声認識部7に、音声認識を開始するように指示を出すものである。ここでは、例えば、話者が（本実施形態の画像認識装置に対して（すなわち画像取得部1の受光素子の部分に対して；以下、同様））正面を向いたとき、話者の行為が開始されたとみなし、この時点で、音声認識の開始の指示を音声認識部7に送る。

【0155】以上のように本実施形態によれば、話者の動作に応じて、音声認識を開始することが可能である。例えば、話者が（本実施形態の画像認識装置に対して）正面を向いたときに音声認識を開始することができる。

【0156】また、本実施形態によれば、画像認識部3による口唇認識（読唇）の結果も得られるため、音声認識と口唇認識（読唇）を同時に行うことが可能となり、これら2つの認識の結果を総合的に用いることにより、話者の会話内容について、より高い認識率を持つ認識結果を得ることができる。

【0157】これは、以下の様な状況において大変有効である。例えば、工事現場などの雑音が多く声を聞き取りにくいような場所では、音声認識だけの場合、認識率が低下するし、場合によっては全然認識できなくなったりするが、第3の実施形態のように、口唇認識も同時に行えば、口唇認識は雑音に影響されないので、認識率は低下することはなく、全体的に高い認識率を維持することができる。また、図書館のような静寂で大きな声を出せない場所でも、音声認識だけでは、微少の音声で認識を行なわねばならないため、認識率の低下が考えられるが、同様の理由で、口唇認識も同時に行えば、全体的に高い認識率を維持することができる。

【0158】また、2人が話をしているような場合、従来の音声認識では、複数の音声が同時に入力されてしまい、認識対象を判別することが困難だったが、本実施形態の場合、2人のうち、例えば、本実施形態の画像認識装置に対して正面を向いている人の方のみを認識するというように、認識対象を判別することも容易であるし、口唇認識も同時に行っているので、その情報を用いて認識対象を判別することもできる。

【0159】本実施形態は、上記した構成に限定されず、種々変形して実施することができる。以下では、本実施形態のいくつかの変形例を示す。

【0160】（第3の実施形態の変形例1）第3の実施形態では、音声認識部7、音声認識開始部8を置き、方向識別部6で得られた結果をもとに、音声認識を開始する例について説明したが、これに限定されるものではなく、音声認識に限らず、他のどのような認識手段でも良い。

【0161】（第3の実施形態の変形例2）第3の実施形態では、話者の顔の向いている向きに応じて、音声認識の開始の指示に用いる例を示したが、図16に示すように、音声認識開始部8の代わりに、音声認識部7に音声認識の終了を指示するための音声認識終了部9を置き、音声認識の終了の指示に用いても良い。

【0162】こうした場合、話者の動作に応じて、音声認識を終了することが可能である。例えば、話者が（本実施形態の画像認識装置に対して）顔を背けたときに音声認識を終了することができる。

【0163】もちろん、図15にさらに音声認識終了部9を設け、音声認識の開始と終了の両方の指示に用いてもよい。

【0164】（第3の実施形態の変形例3）方向識別部6で話者の顔の向いている方向を得て、それを音声認識の開始の指示に用いるのではなく、図17に示すように、画像認識部3で得られた認識結果から、会話の始まりにおける口唇の動き出しを検出し、それをもとに音声認識部7に音声認識の開始を指示するための新たな音声認識開始部8を置いても良い。

【0165】この場合、音声認識開始部8は、画像認識部3で得られた口唇認識の結果から、口唇の動作が始ま

る点（言葉を話し始める際、口唇が微妙に動き始める点で、この時点では、まだ発音は始まっていない）を求め、その時点で、音声認識部７に音声認識の開始を指示する。

【０１６６】また、同様に、本変形例３の音声認識開始部８の代わりに、口唇の動作が終了する点を検出する音声認識終了部９を置き、音声認識の終了の指示に用いても良い。

【０１６７】もちろん、同様に、本変形例３の音声認識開始部８に加えて、口唇の動作が終了する点を検出する音声認識終了部９を置き、音声認識の開始と終了の両方の指示に用いても良い。

【０１６８】なお、従来方法では、口唇の動きだしの検出をおこなうための計算に時間がかかるため、このようなリアルタイム処理に口唇の動きだしの検出を用いることは困難であったが、本実施形態の画像認識装置では、第１の実施形態で説明したように、あまり計算コストを必要とせずに口唇部の抽出が可能であるため、このような口唇の動きだしの検出を十分にリアルタイムに行うことができる。

【０１６９】（第４の実施形態）次に、本発明の第４の実施形態について説明する。本実施形態では、第１の実施形態と相違する部分を中心に説明する。

【０１７０】図１８は、本発明の第４の実施形態に係る画像認識装置の全体構成図である。

【０１７１】図１８に示されるように、本実施形態の画像認識装置は、第２の実施形態の画像認識装置の構成に対して、各種の情報の提示を行う情報呈示部１０と、方向識別部６で得られた話者の顔の向いている方向をもとに情報呈示の開始を情報呈示部１０に指示するための情報呈示開始部１１が追加された構成になっている。

【０１７２】これにより話者の顔の向いている方向に応じて、各種の情報呈示を行うことができる。

【０１７３】次に、情報呈示部１０について説明する。

【０１７４】情報呈示部１０は、対象者（話者）に何らかの情報を提示するものである。情報呈示部１０は、ディスプレー（画像、文字などを呈示）、スピーカー（音を呈示）、フォースフィードバック装置（感触を呈示）などの少なくとも１つの情報呈示装置を具備しており、それを通して対象者に情報を提示することができる。

【０１７５】次に、情報呈示開始部１１について説明する。

【０１７６】情報呈示開始部１１は、前述した第３の実施形態における音声認識開始部８と同様の役割をするもので、方向識別部６で得られた結果をもとに、情報呈示部１０に、情報呈示の開始の指示を出すものである。

【０１７７】本実施形態によれば、話者の動作に応じて、情報呈示を開始することが可能である。例えば、話者が（本実施形態の画像認識装置に対して）正面を向いたときに、それを話者の行為開始とみなし、情報呈示を開始することができる。

【０１７８】また、画像認識部３による口唇認識（読唇）の結果も得られているため、話者の会話の内容に応じて、情報呈示を開始することも可能である。

【０１７９】本実施形態は、上記した構成に限定されず、種々変形して実施することができる。以下では、本実施形態のいくつかの変形例を示す。

【０１８０】（第４の実施形態の変形例１）第３の実施形態の変形例２の場合と同様に、情報呈示開始部１１に代えてあるいは情報呈示開始部１１に加えて、情報呈示終了部を置き、呈示終了の指示をしても良い。

【０１８１】（第４の実施形態の変形例２）第３の実施形態の変形例３の場合と同様に、画像認識部３で得られた認識結果から、会話の始まりにおける口唇の動き出しを検出し、それをもとに情報呈示部１０に情報呈示の開始を指示するための新たな情報呈示開始部１１を置いても良い。

【０１８２】このようにすることにより、例えば、情報呈示の方法として音声合成を用いて、口唇の形状、動きの認識結果をもとに、その認識内容を音声合成で提供することで、喉の病気などで言葉が話せない場合でも、口パク（音声は出さずに、実際話しているように口唇を動かす）をするだけで、音声合成により、本実施形態の画像認識装置に代わりに話させるなどというような、いわゆる、音声同期（リップシンク）が可能である。

【０１８３】もちろん、第３の実施形態の変形例３の場合と同様に、本変形例の情報呈示開始部１１に代えてあるいは情報呈示開始部１１に加えて、情報呈示終了部を置き、呈示終了の指示をしても良い。

【０１８４】（第４の実施形態の変形例３）図１９に示すように、情報呈示開始部１１の代わりに、呈示する情報の種類を切り替えるための情報呈示切り替え部１２を置き、話者の向いている方向によって、情報呈示の形態を切り替えるようにしても良い。

【０１８５】この情報呈示の形態の切り替えとしては、（１）異なる情報呈示の形態を追加する、（２）複数の情報呈示の形態を提供している場合に、少なくとも１つの情報呈示の形態を中止する、（３）１または複数の情報呈示の形態を提供している場合に、一部または全てを異なる情報呈示の形態に変更する（情報呈示の形態数が変化する場合を含む）、などが考えられる。

【０１８６】こうすることで、話者の顔が（本実施形態の画像認識装置の方を）向いていないときには、音声のみの情報呈示を行っていて、話者の顔が向いたときには、情報呈示切り替え部１２を用いて、音声のみの呈示から、音声に加えて、画像などの複合メディアを用いた情報呈示に切り替える、などということが可能である。

【０１８７】これは、例えば、博物館、美術館などの展示物の説明を行うのに、通常は音声で説明文を読み上げておいて、見学者が展示物の方を見て（あるいは、さら

に何か話すと）、展示物の横に置いておいたディスプレーで説明ビデオの上映が始まる、といったように用いることができる。

【0188】（第4の実施形態の変形例4）第4の実施形態に、第3の実施形態で説明した音声認識部、音声認識開始部、音声認識終了部などを組み合わせることにより、話者の生の音声と情報呈示部10で生成した画像情報を組み合わせて呈示することが可能となる。

【0189】例えば、口腔部抽出部2で抽出した口腔部の距離画像ストリームを用いて、情報呈示部10でその形状を3次元CG合成を行い、それに、音声認識部で取得した話者の生の音声を組み合わせることで、話者の生の声と音声同期（リップシンク）して口唇が動く3次元CGを提供することができる。

【0190】（第5の実施形態）次に、本発明の第5の実施形態について説明する。

【0191】第5の実施形態の画像認識装置は、第1、第2、第3、あるいは第4の実施形態の画像認識装置やそれらの種々の変形例の構成それぞれにおいて、外部との通信を行う通信部（図示せず）を追加したものである。

【0192】これにより第1、第2、第3、あるいは第4の実施形態やその変形例で得られた所望の情報を外部に通信することができる。

【0193】通信部は、入力されたデータを、電話回線などの通信路を用いて外部に通信するもので、これが加えられることで、例えば、第1の実施形態では、口唇認識の結果を、第2の実施形態では、口唇認識の結果および話者の向いている方向を、第3の実施形態では、口唇認識および音声認識の結果を、第4の実施形態では、口唇認識の結果および呈示された情報を、それぞれ通信することが可能である。

【0194】以上のように本実施形態によれば、当該画像認識装置で得られた結果（第1の実施形態を基にしたものでは、口唇認識結果、第2の実施形態を基にしたものでは、話者方向と口唇認識結果、第3の実施形態を基にしたものでは、口唇および音声認識結果、第4の実施形態を基にしたものでは、口唇認識結果および呈示情報）を、インターネットなどを通して通信することが可能である。

【0195】例えば、第4の実施形態の変形例4の場合、話者の生の声と音声同期（リップシンク）して口唇が動く3次元CGが得られるが、先に顔の口唇部以外の部分を通信先の相手に送っておき、話者の発言とともに、上記3次元CGの口唇部だけを通信部を用いてリアルタイムに送り、通信先で、あらかじめ送っておいた顔と合成することで、通信路に負荷をかけずに（つまり通信路をボトルネックとせずに）、3次元CGの音声同期（リップシンク）を行うことができる。これは、通信路に速度のボトルネックが生じやすいインターネットなどで、音声とCGといった比較的大きなデータを用いてリアルタイム処理する際に大変有効である。

【0196】以下では、以上の各実施形態における画像取得部1の構成について詳しく説明する。

【0197】図20に、画像取得部1の一構成例を示す。この画像取得部1は、対象物体に光を照射するための発光部101、対象物体からの反射光を画像として抽出するための反射光抽出部102、画像化された反射光の情報をもとに距離画像を生成するための距離画像生成部103、これらの各部の動作タイミングを制御するタイミング制御部104を用いて構成される。

【0198】発光部101は、発光素子を持ち、タイミング制御部104によって生成されるタイミング信号に従って時間的に強度変動する光を発光する。発光部101が発した光は、発光部101の発光素子の前方にある対象物体により反射された後に、反射光抽出部102の受光面に入射する。

【0199】物体からの反射光は、物体の距離が大きくなるにつれ大幅に減少する。物体の表面が一様に光を散乱する場合、反射光画像1画素あたりの受光量は物体までの距離の2乗に反比例して小さくなる。従って、当該受光面の前に物体が存在する場合、背景からの反射光はほぼ無視できるくらいに小さくなり、物体のみからの反射光画像を得ることができる。

【0200】例えば、当該受光面の前に人間の顔の部分が存在する場合、その顔からの反射光画像が得られる。このとき、反射光画像の各画素値は、その画素に対応する単位受光部で受光した反射光の量を表す。反射光量は、物体の性質（光を鏡面反射する、散乱する、吸収する、など）、物体の向き、物体の距離、などに影響されるが、物体全体が一様に光を散乱する物体である場合、その反射光量は物体までの距離と密接な関係を持つ。顔などはこのような性質を持つため、顔を対象物体とした場合の反射光画像は、顔の3次元形状、顔の距離、顔の傾き（部分的に距離が異なる）、などを反映する。

【0201】反射光抽出部102は、マトリクス状に配列した、光の量を検出する受光素子を持ち、発光部101が発した光の対象物体による反射光の空間的な強度分布を抽出する。この反射光の空間的な強度分布は、画像として捉えることができるので、以下では反射光画像と呼ぶ。

【0202】ここで、反射光抽出部102の受光素子においては、一般的に、発光部101の光の対象物体による反射光だけでなく、照明光や太陽光などの外光も同時に受光することが想定される。そこで、本構成例の反射光抽出部102では、発光部101が発光しているときに受光した光の量と、発光部101が発光していないときに受光した光の量の差を取ることによって、発光部101からの光の対象物体による反射光の成分だけを取り出すようにしている。この受光のタイミングも、タイミ

ング制御部１０４によって制御される。

【０２０３】そして、反射光抽出部１０２により得られた外光補正後の反射光画像の各画素に対応する反射光量（アナログ信号）が必要に応じて増幅された後にA／D変換され、これによってデジタル化された反射光画像が得られる。

【０２０４】距離画像生成部１０３は、反射光抽出部１０２によって得られた反射光画像の各画素の受光量の値（デジタルデータ）を距離の値に変換することによって、距離画像（例えば、６４画素×６４画素、２５６階調の画像）を生成する。

【０２０５】次に、図２１に、画像取得部１のより詳しい一構成例を示す。

【０２０６】発光部１０１より発光された光は、対象物体１０６に反射して、レンズ等の受光光学系１０７により、反射光抽出部１０２の受光面上に結像する。

【０２０７】反射光抽出部１０２は、この反射光の強度分布、すなわち反射光画像を検出する。反射光抽出部１０２は、各画素（単位受光部）ごとに設けられた第１の受光部１２１および第２の受光部１２２、ならびに全画素について１つ（または一纏まりの複数画素ごとにまたは各画素ごとに）設けられた差分演算部１２３を用いて構成される。

【０２０８】第１の受光部１２１と第２の受光部１２２は、異なるタイミングで受光を行う。そして、第１の受光部１２１が受光しているときに発光部１０１が発光し、第２の受光部１２２が受光しているときには発光部１０１は発光しないように、タイミング制御部１０４がこれらの動作タイミングを制御する。これにより、第１の受光部１２１が発光部１０１からの光の物体による反射光とそれ以外の太陽光、照明光などの外光を受光する。一方、第２の受光部１２２は外光のみを受光する。両者が受光するタイミングは異なっているが近いので、この間における外光の変動や対象物体の変位は無視できる。

【０２０９】従って、差分演算部１２３により第１の受光部１２１で受光した像と第２の受光部１２２で受光した像の差分をとれば、対象物体による反射光の成分だけが抽出される。１つの差分演算部１２３が複数の画素で共用される場合には、シーケンシャルに差分が演算される。

【０２１０】なお、単位受光部の第１の受光部１２１および第２の受光部１２２の実際の構成については種々のものが考えられる。例えば、第１の受光部１２１および第２の受光部１２２のそれぞれに受光素子を設けるのではなく、単位受光部ごとに、光電変換素子（例えばフォトダイオード）を１つ設けて第１の受光部１２１と第２の受光部１２２で兼用するとともに、受光量に対応する電荷量を蓄積する電荷蓄積素子（例えばコンデンサ）を第１の受光部１２１および第２の受光部１２２のそれぞれのために２つ設ける方法が考えられる。

【０２１１】さて、上記のようにして、反射光抽出部１０２は、反射光画像の各画素の反射光量を外光補正を行った後に出力する。なお、ここでは、各画素の反射光量をシーケンシャルに出力するものとする。

【０２１２】反射抽出部１０２からの出力はアンプ１３１によって増幅され、A／D変換器１３２によってデジタルデータに変換された後、メモリ１３３に画像データとして蓄えられる。そして、しかるべきタイミングでこのメモリより蓄積されたデータが読み出され、距離画像生成部１０３に与えられる。

【０２１３】距離画像生成部１０３では、反射光抽出部１０２により得られた反射光画像をもとに距離画像を生成する。例えば、反射光画像の各画素の反射光量を、それぞれ、所定の階調（例えば、２５６階調）のデジタルデータに変換する。なお、この変換にあたっては、例えば、（１）受光素子における受光量が対象物体までの距離に対して非線形性を持つ（対象物体までの距離の２乗に反比例する）という非線形要因に対する補正を行う処理、あるいは（２）各画素に対応する受光素子の特性のばらつきや非線形性を補正する処理、あるいは（３）背景やノイズを除去する処理（例えば、基準値以下の受光量を持つ画素の階調を０にする）、などといった処理を適宜行ってもよい。

【０２１４】なお、顔の３次元形状を抽出する場合、距離情報を高い分解能で求められることが望ましい。この場合、アンプ１３１として対数アンプを用いると望ましい。受光面での受光量は対象物体までの距離の２乗に反比例するが、対数アンプを用いると、その出力は距離に反比例するようになる。このようにすることで、ダイナミックレンジを有効に使うことができる。

【０２１５】さて、上記のような構成において、１回の発光によって全画素について反射光が得られるものとすると、タイミング制御部１０４の制御によって、発光→第１の受光部による受光→発光なしに第２の受光部による受光→差分演算→デジタル化→距離画像の生成（もしくは発光なしに第２の受光部による受光→発光→第１の受光部による受光→差分演算→デジタル化→距離画像の生成）といった一連の処理が進められ、これによって１枚の距離画像が得られる。また、この一連の処理を繰り返し行う（例えば、１／６０秒ごとに行う）ことによって、距離画像ストリームを得ることができる。

【０２１６】なお、発光部１０１は、人間の目に見えない、近赤外光を発光するようにするのが好ましい。このようにすれば、光が照射されても人間には光が見えないため、眩しさを感じさせないようにすることができる。また、この場合に、受光光学系には、近赤外光通過フィルタを設けると好ましい。このフィルタは、発光波長である近赤外光を通過し、可視光、遠赤外光を遮断するので、外光の多くをカットすることができる。ただし、人

間の目に眩しくない条件であれば（例えば、発光量がそれほど大きくない、人間の目には直接入射しないような光学系となっている、など）、可視光を用いても構わない。また、電磁波や超音波などを用いる方法も考えられる。

【0217】また、上記では外光補正として発光部101の発光の有無の相違による2種類の受光量の差分をアナログ信号の状態で取ったが、2種類の受光量をそれぞれデジタル化した後に差分を取るようにする方法もある。

【0218】なお、上記した受光面もしくはこれを収容した筐体は、本画像認識装置の目的等に応じて適宜設置するばよい。例えば本画像認識装置が表示装置を持つものである場合、この表示装置に対して対象物体となる人間の顔が正面を向いたときに、当該受光面に対しても正面を向いた形になるように当該画像認識装置の筐体に設ける。

【0219】なお、以上の各実施形態やその変形例は、適宜組み合わせて実施することが可能である。

【0220】また、以上の各実施形態やその変形例あるいはそれらを適宜組み合わせたものでは、距離画像ストリームから形状および／または動きを認識し、あるいはさらにその認識結果をもとに種々の処理を行うものであったが、距離画像から形状を認識し、あるいはさらにその認識結果をもとに種々の処理を行うように構成した実施形態も可能である。

【0221】また、以上の各実施形態やその変形例あるいはそれらを適宜組み合わせたものは、画像取得部1もしくはそのうちの反射光画像を抽出する部分を省き、与えられた距離画像もしくはそのストリームに基づいて、もしくは与えられた反射光画像もしくはそのストリームから距離画像もしくはそのストリームを生成し、生成した距離画像もしくはそのストリームに基づいて、形状および／または動きを認識し、あるいはさらにその認識結果をもとに種々の処理を行うような装置として構成することも可能である。

【0222】以上の各機能は、素子部分を除いて、ソフトウェアとしても実現可能である。また、上記した各手順あるいは手段をコンピュータに実行させるためのプログラムを記録した機械読取り可能な媒体として実施することもできる。

【0223】本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0224】

【発明の効果】本発明によれば、対象物体に対する距離画像から必要とする部分を抽出し、抽出した部分の距離画像に基づいて認識処理を行うので、人間の顔や口唇の形状や動きを高速かつ高精度に認識することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係る画像認識装置の構成例を概略的に示す図

【図2】距離画像について説明するための図

【図3】距離画像について説明するための図

【図4】距離画像について説明するための図

【図5】エッジ抽出の処理の流れを示すフローチャート

【図6】Ｓｏｂｅｌオペレータを説明するための図

【図7】テンプレートマッチングの処理の流れを示すフローチャート

【図8】本発明の第1の実施形態の変形例2に係る画像認識装置の構成例を概略的に示す図

【図9】本発明の第1の実施形態の変形例3に係る画像認識装置の構成例を概略的に示す図

【図10】本発明の第2の実施形態に係る画像認識装置の構成例を概略的に示す図

【図11】話者の顔の向いている方向を求める処理の流れを示すフローチャート

【図12】画素の法線方向を説明するための図

【図13】本発明の第2の実施形態の変形例1に係る画像認識装置の構成例を概略的に示す図

【図14】本発明の第2の実施形態の変形例2に係る画像認識装置の構成例を概略的に示す図

【図15】本発明の第3の実施形態に係る画像認識装置の構成例を概略的に示す図

【図16】本発明の第3の実施形態の変形例2に係る画像認識装置の構成例を概略的に示す図

【図17】本発明の第3の実施形態の変形例3に係る画像認識装置の構成例を概略的に示す図

【図18】本発明の第4の実施形態に係る画像認識装置の構成例を概略的に示す図

【図19】本発明の第4の実施形態の変形例3に係る画像認識装置の構成例を概略的に示す図

【図20】画像取得部の構成例を示す図

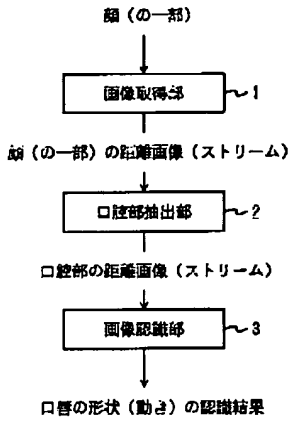
【図21】画像取得部のより詳しい構成例を示す図

【符号の説明】

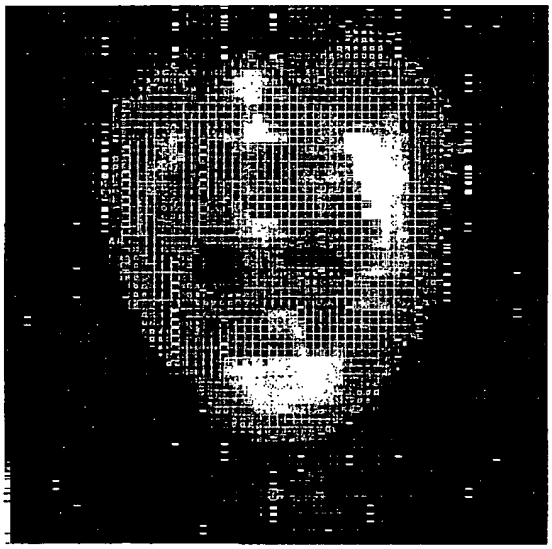
1…画像取得部
2…口腔部抽出部
3…画像認識部
4…音呈示部
5…顔部抽出部
6…方向識別部
7…音声認識部
8…音声認識開始部
9…音声認識終了部
10…情報呈示部
11…情報呈示開始部
12…情報呈示切り替え部
101…発光部
102…反射光抽出部
103…距離画像生成部

１０４…タイミング信号生成部
１０７…受光光学系
１２１…第１の受光部
１２２…第２の受光部
１２３…差分演算部
１３１…アンプ１３１
１３２…Ａ／Ｄ変換器
１３３…メモリ

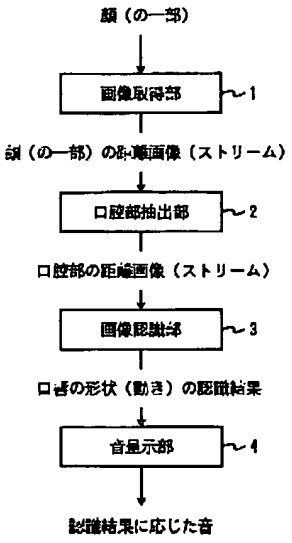
【図１】



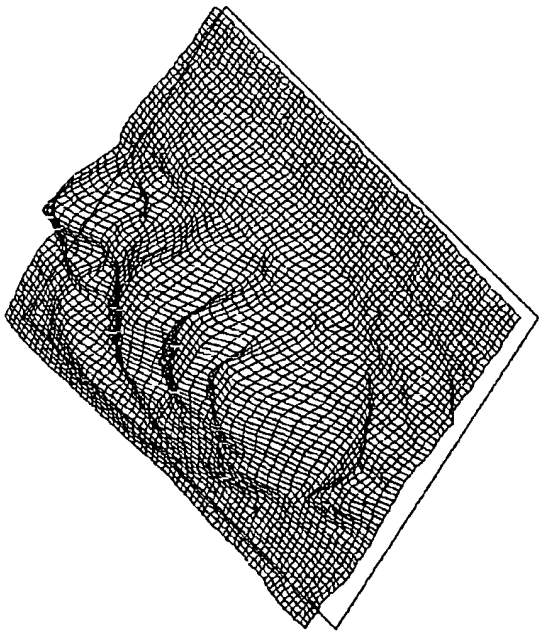
【図２】



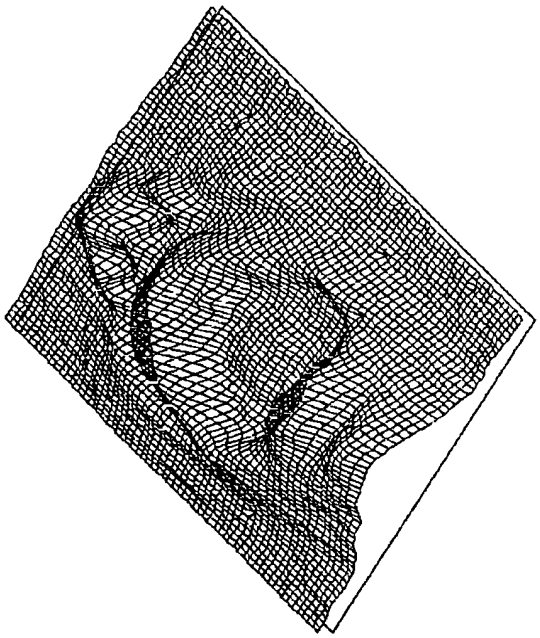
【図８】



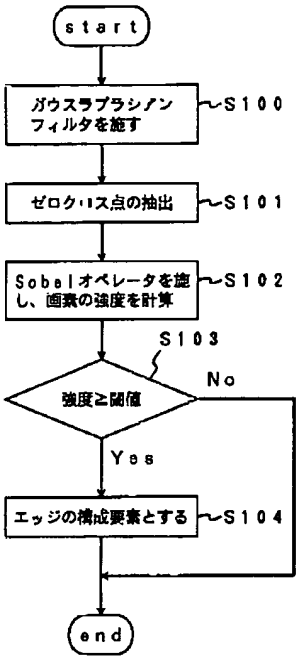
【図３】



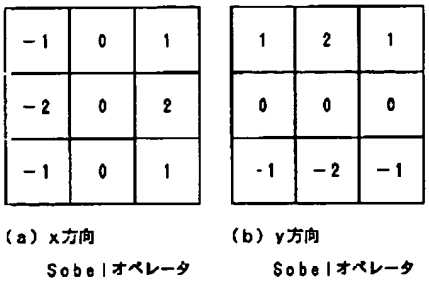
【図４】



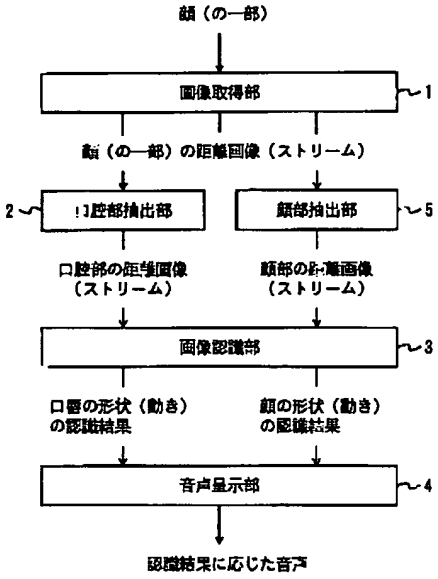
【図5】



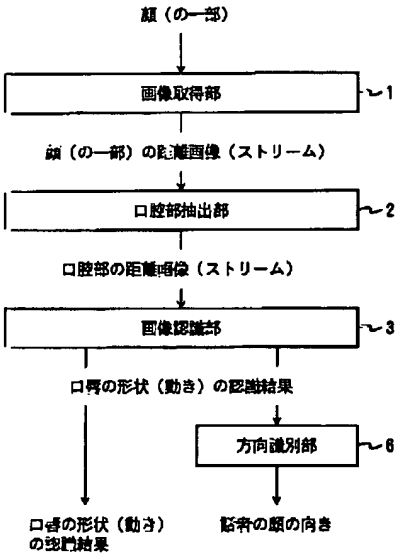
【図6】



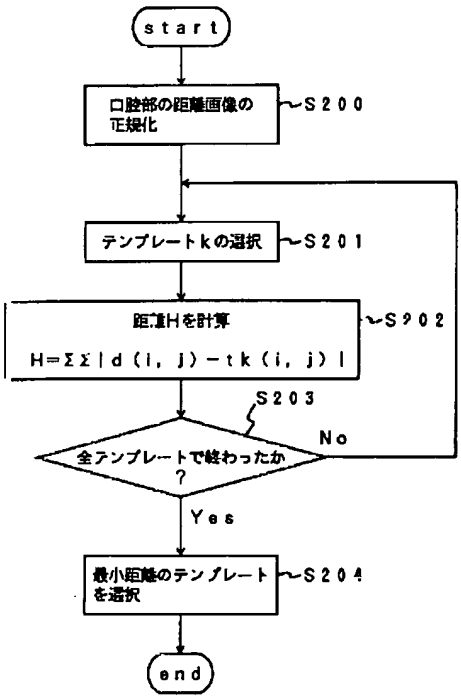
【図9】



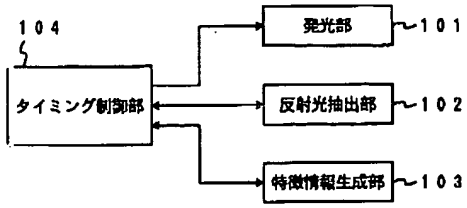
【図10】



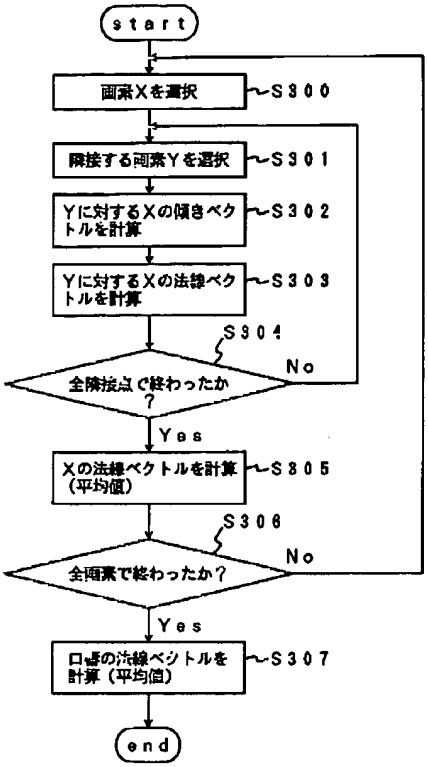
【図7】



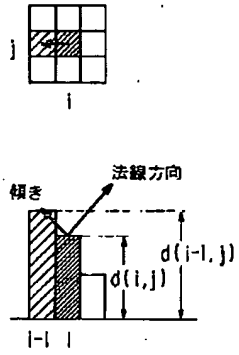
【図20】



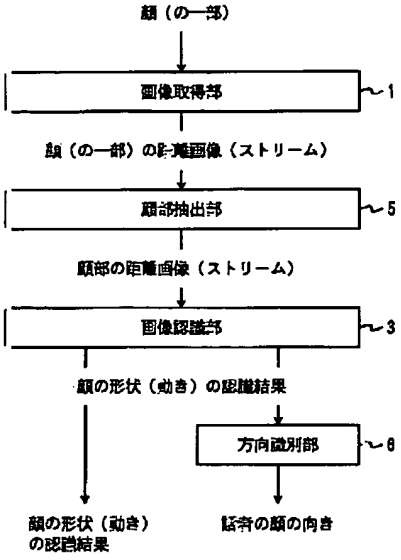
【図11】



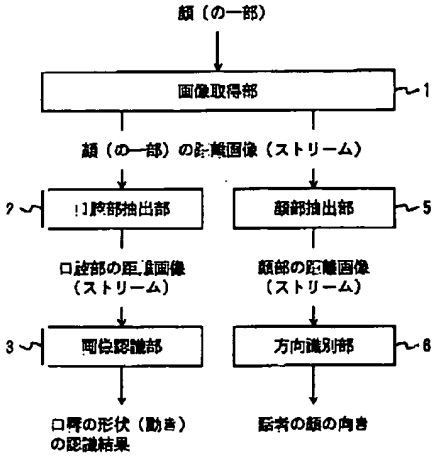
【図12】



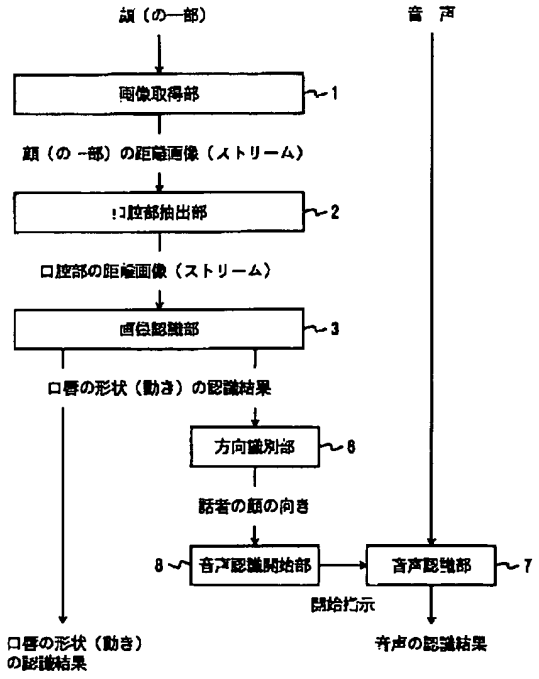
【図13】



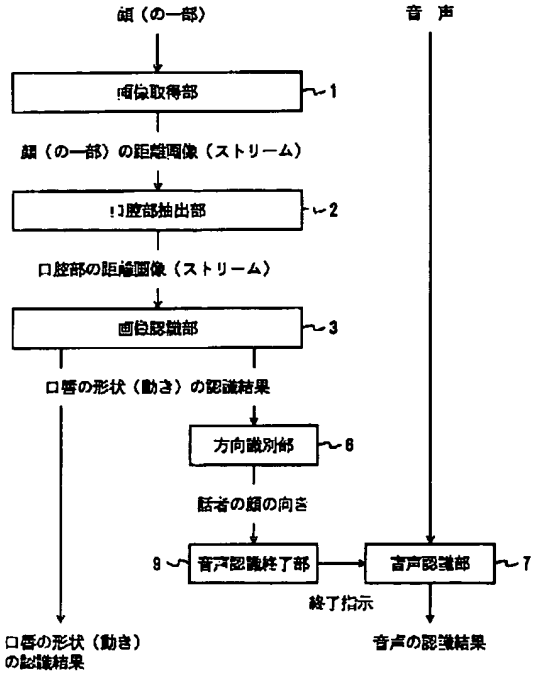
【図14】



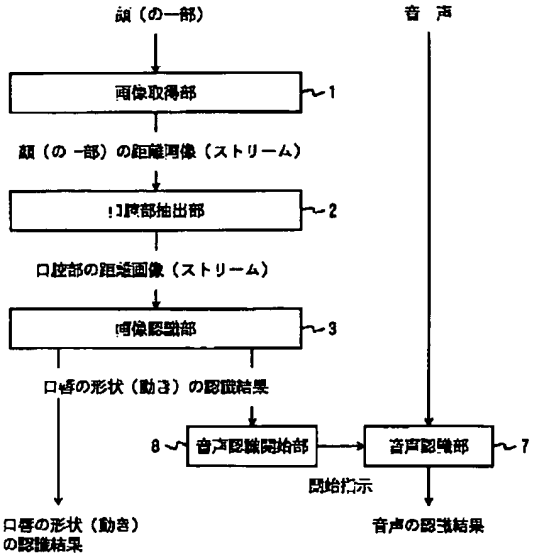
【図15】



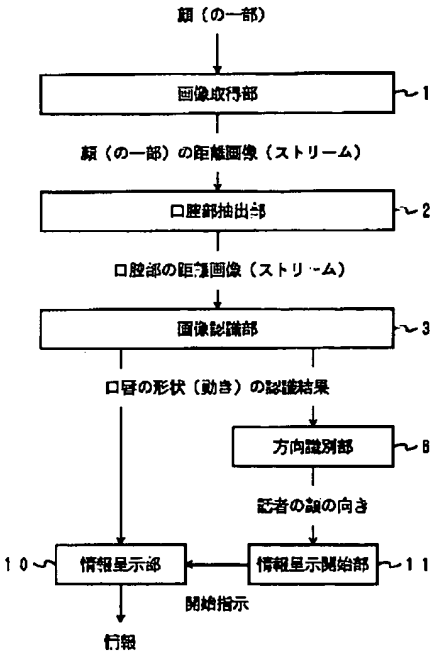
【図16】



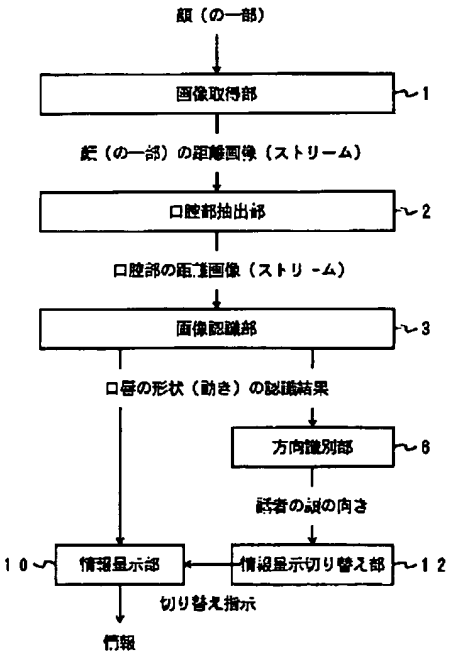
【図17】



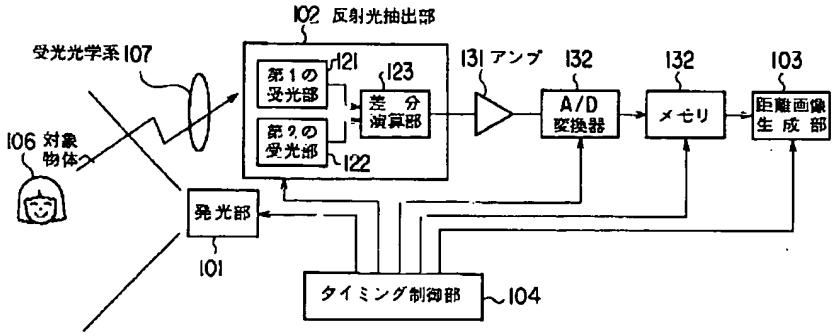
【図18】



【図19】



【図21】


102 反射光抽出部
受光光学系 107
第1の受光部
121
123
差分演算部
131 アンプ
132
A/D変換器
132
メモリ
103
距離画像生成部
106 対象物体
第2の受光部
122
発光部
101
タイミング制御部
104

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.